

## SYSTEM AND METHOD FOR PARALLEL CONVERSION, COMPILATION, AND DISTRIBUTION OF CONTENT

### FIELD OF THE INVENTION

[0001] The present invention relates to the field of providing content, and, more specifically, to converting content into different formats and compiling and distributing the converted content.

### BACKGROUND OF THE INVENTION

[0002] Content providers define content in a wide variety of formats. For example, extensible markup language (XML) is a commonly used format that enables designers to create customized tags, thereby enabling the transmission, validation, and interpretation of data between applications. Content defined in XML includes a hierarchy of elements and attributes. For example, a test defined in XML may include a highest level element specifying a type of test such as, for example, a mathematics test. A next level element may specify a section of the test such as, for example, geometry. A next level element may specify an individual question, and a next level element may specify elements of the test question such as, for example, a question text (“stem”), the correct answer (“key”), and a set of possible answers (“distracters”).

[0003] XML is particularly suited for defining content for a variety of transmissions, including the Internet. For example, the mathematics test discussed above may be distributed from a server to a “presentation” client at a test presentation center via the Internet. A user interface such as, for example, a web browser running at the presentation client may be used to request, receive, and present the test. The test may be displayed to a test taker on a display device such as, for example, a monitor attached to the presentation client. The test taker may interact with the test using input devices such as, for example, a keyboard and a mouse. Feedback related to presentation of the test such as, for example, test answers and personal information about the test taker may be submitted from the presentation client back to the server.

[0004] In addition to presenting content using a web browser, it is often desirable to present content using alternative presentation mediums. For example, it may be desirable to

present a standard printed or “published” version of the mathematics test. Furthermore, it may be desirable to present an audible version of the test using technology such as, for example, interactive voice response (IVR). Consequently, it may be desirable to convert content from a format such as XML to alternative formats that are specific to the alternative presentation mediums. For example, portable document format (PDF) is a file format developed by Adobe Systems of San Jose, California, that is well suited for distributing and presenting a printed version of content. The “WAV” file format is a file format well known in the art of content provision that is well suited for presenting audible versions of content.

[0005] In addition to converting content into such alternative formats, it is often also desirable to compile such converted content so that it may be distributed as a single compilation. For example, it may be desirable to compile a published PDF version and a WAV version of the mathematics test and to distribute the single compilation to multiple presentation clients. The distribution of such a compilation may enable, for example, an audible version of a test to be presented to visually impaired test takers and a printed version of a test to be delivered to hearing impaired test takers. Furthermore, for example, the converted versions of the mathematics test may be bundled with an XML file that specifies a list of authorized test delivery centers to which the mathematics test should be distributed. Additionally, it may also be desirable to bundle another XML file that specifies a list of registered test takers at each such authorized delivery center.

[0006] Thus, there is a need for a processing and publishing “pipeline” for parallel conversion, compilation, and distribution of content. It is desired, for example, that such a pipeline provide the application framework and interface to enable a user to request that a plurality of content items be converted, in parallel, to a plurality of different formats. It is further desired, for example, that the application framework and interface to enable a user to request the compilation and distribution of such converted content items.

## **SUMMARY OF THE INVENTION**

[0007] An application framework and interface may be provided that enables a user to request the conversion of content items into formats well suited for presentation. The request may include a set of deliverables. Each deliverable may include an associated content item and a corresponding associated format to which to convert the associated content item. The request

may also include selected presentation devices or delivery channels to which to distribute the converted content items. The request may also include customization options related to the presentation of the content items at the selected presentation devices or delivery mechanisms.

[0008] The request may be processed by a “publishing pipeline” application. The pipeline application may log the request in a request history and retrieve the content items from a database. The pipeline application may prepare the content items for conversion into the corresponding formats. Such preparation may include, for example, retrieving the content items from a content database and customizing the content items. The pipeline application may convert the associated content items using a plurality of parallel processing threads, each thread corresponding to an associated deliverable, whereby each thread converts the associated content item to the corresponding associated format. The pipeline application may compile the converted content, and distribute the compiled content to the selected presentation devices over a selected delivery channel.

[0009] The distributed content may be received and presented at the selected presentation devices or delivery channel. A printer attached to a presentation device may be used, for example, to present a printed version of content. A speaker attached to a presentation device may be used, for example, to present an audible version of content. Feedback may be generated based on the presentation of the content.

[0010] Additional features and advantages of the invention will be made apparent from the following detailed description of illustrative embodiments that proceeds with reference to the accompanying drawings.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

[0011] The illustrative embodiments will be better understood after reading the following detailed description with reference to the appended drawings, in which:

[0012] Fig. 1 is a block diagram of an exemplary content distribution and presentation system in accordance with the present invention;

[0013] Fig. 2 is a block diagram of an exemplary content conversion, compilation, and distribution system in accordance with the present invention;

[0014] Fig. 3 is an exemplary user interface for requesting conversion, compilation, and distribution of content in accordance with the present invention;

- [0015] Fig. 4 is an exemplary processing pipeline for content conversion, compilation, and distribution in accordance with the present invention;
- [0016] Fig. 5 is a cross sectional view of the content conversion portion of the pipeline of Fig. 4 in accordance with the present invention; and
- [0017] Fig. 6 is flowchart of an exemplary method for content conversion in accordance with the present invention.

## **DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS**

[0018] Systems and methods that meet the above-mentioned objects and provide other beneficial features in accordance with the presently preferred exemplary embodiments of the invention will be described below with reference to aforementioned Figures. Those skilled in the art will readily appreciate that the description given herein with respect to those figures is for explanatory purposes only and is not intended in any way to limit the scope of the invention. Throughout the description, like reference numerals will refer to like elements in the respective figures.

### **Exemplary Conventional Content Distribution and Presentation System**

[0019] An exemplary conventional content distribution and presentation system 100 in accordance with the present invention is shown in Fig. 1. Generally, content items may be stored at content database 110. Such content items may be retrieved by server 120 via communications network 140 and distributed to presentation clients 130a-n via communications network 140. The distributed content items may be received and presented at presentation clients 130a-n.

[0020] Content database 110 may be, for example, a relational database management system (RDBMS) that enables the storage and manipulation of content. As should be appreciated, multiple databases or other content storage sources may be used in accordance with the present invention. Content may include items such as text, video, and audio items. Specifically, for example, such content items may include articles, novels, tests, movies, songs, and reports.

[0021] Presentation clients 130a-n and server 120 may be computing devices such as, for example, personal computers (PC's). Communications network 140 may be a fixed-wire or wireless network which may be a local area network (LAN), wide area network (WAN), intranet,

the Internet, or other computer network. If communications network 140 is the Internet, for example, then server 120 can be a Web server with which the presentation clients 130a-n may communicate via any of a number of known communication protocols, such as, hypertext transfer protocol (HTTP), simple mail transfer protocol (SMTP) or file transfer protocol (FTP).

[0022] Presentation clients 130a-n can be equipped with a Web browser to gain access to the server 120. Presentation clients 130a-n may be located at presentation facilities and may be used in connection with attached presentation devices such as, for example, monitors, speakers, or printers. Such monitors may be used, for example, for presenting computer based tests. Such speakers may be used, for example, for displaying computer based versions of content. Such printers may be used, for example, for presenting “published” or printed versions of content. Presentation clients 130a-n may also be used in connection with attached input devices such as, for example, keyboards, mice, or microphones, which enable interaction with the presented content. Feedback related to presentation of the content may be submitted from presentation clients 13-a-n back to server 120 via communication network 140.

#### **Systems and Methods for Content Retrieval, Conversion, Compilation, and Distribution**

[0023] An exemplary content conversion, compilation, and distribution system 200 in accordance with the present invention is shown in Fig. 2. Generally, request client 210 provides a user interface 210a for a participating user to request parallel conversion, compilation, and distribution of content. Such a request may be submitted to server 120 via communications network 140. Pipeline processing application 122 running at server 120 may process the request by retrieving content from content database 110, converting the content to selected file formats in parallel, compiling the converted content, and distributing the compiled content. Plug-ins 220a-n may be used in connection with server 120 to convert content to selected file formats.

[0024] Request client 210 may be a computing device such as, for example, a PC. As should be appreciated, request client 210 may be one of presentation clients 130a-n of Fig. 1. User interface 210a may be an application running at request client 210. User interface 210a may be a software application enabling communication with server 120. If, for example, communications network 140 is the Internet, then user interface 210a may be a Web browser. If, for another example, the communications network is a LAN, then the user interface 210a may be another application that is making the request. As should be appreciated, communications between request client 210 and server 120 may be embedded communications.

[0025] User interface 210a generally enables a participating user to request content items to be converted in parallel, compiled, and distributed. An exemplary user interface 210a in accordance with the present invention is shown in Fig. 3. As should be appreciated, although exemplary user interface 210a is described with respect to test content, a user interface in accordance with the present invention is not limited to use in connection with test content and may be used in connection with any type of content. User interface 210a may be presented via a display device such as, for example, a monitor attached to request client 110. The user may be authorized to access user interface 300 based on, for example, a valid password or identifier.

[0026] Select content item drop down menu 310 enables the user to select content items to be retrieved, converted, compiled, and distributed. As should be appreciated, select content item drop down menu 310 and other drop down menus 320 and 340 need not necessarily be drop down menus and may, for example, be traditional menu sets including folders and documents, lists of hyperlinks, or separate windows. Content item menu 310 includes exemplary content items such as “Mathematics Test” and “English Test”. The content items included in content item menu 310 may be stored at content database 110. Such content items may be defined in a format such as, for example, extensible markup language (XML), that is well suited for transmitting and presenting content over the Internet or other delivery channels. A separate authorization may be required for the user to request conversion and distribution of each individual content item shown in content item menu 310.

[0027] Select conversion format drop down menu 320 enables the user to select formats to which to convert menu items selected from content item menu 310. Conversion format menu 320 includes the exemplary conversion formats portable document format (PDF) and “WAV”. PDF is a file format developed by Adobe Systems of San Jose, California, that is well suited for distributing and presenting a printed version of content. The “WAV” file format is a file format well known in the art of content provision that is well suited for presenting audible versions of content. For example, if a printed version of the “Mathematics Test” is desired, the “Mathematics Test” may be converted from XML to PDF. If an audible version of the “Mathematics Test” is also desired, then “Mathematics Test” may be also be converted from XML to WAV. As should be appreciated, the present invention is not limited to conversion into PDF and WAV file formats. Allowing for the conversion of XML to any delivery channel specific formats is contemplated in accordance with the present invention.

[0028] Plug-ins 220a-n of Fig. 2 may be used in accordance with the present invention in connection with server 120 to parse and convert content to the file format selected in conversion format menu 320 or another selected file format. Plug-ins 220a-n may be commercially available pug-ins, other applications, or proprietary. For example, XML pull parser (XPP) is a plug in that is well known in the art of content parsing that may be used to parse XML content and convert the parsed content to other formats such as, for example, PostScript, which is a format well known in the art developed by Adobe Systems of San Jose, California. An XPP plug-in may, for example, parse an XML file for elements and attributes necessary to convert the content from XML to another selected format.

[0029] View selected deliverables field 330 enables the user to view content that he has selected to be retrieved and converted. Each deliverable includes a content item selected from content item menu 310 and a conversion format selected from conversion format menu 320. View selected deliverables field 330 shows that the “Mathematics Test” and the “English Test” have been selected to be converted from an XML file into both a PDF file and a WAV file.

[0030] Select delivery channels drop down menu 340 enables the user to a delivery channel over which to distribute the deliverables shown in view deliverables field 330. Select delivery channels drop down menu 340 includes exemplary delivery channels “1”, “2”, and “3”. “Select another deliverable” prompt 350a enables the user to select another deliverable to add to view selected deliverables field 330. “Complete request” prompt 350b enables the user to complete and submit his request.

[0031] User interface 210a may engage any application that possesses content that needs to be converted and distributed via a specific delivery channel. Furthermore, user interface 210a is not limited to the menus and fields 310-340 shown in Fig 3 and other additional features, menus, and configuration options are contemplated in accordance with the present invention. For example, “customization” menus may be included which enable a user to specify a time window in which content or portions of content may be presented at presentation clients 130a-n. The time window may me defined by a start timestamp and a stop timestamp. The timestamp may be a value maintained by server 120. Such a time window is a security feature which may prevent unauthorized presentation of the content at an improper time. Another customization menu, for example, may enable the user to select a language such as, for example, Spanish or French, to which to translate content. Another customization option may enable, for example, a presentation authorization. Such a presentation authorization may be, for example, a valid

password or identifier that is required to access and present each deliverable at each presentation client 130a-n.

[0032] An exemplary method for content conversion, compilation, and distribution in accordance with the present invention is shown in Fig. 4. Generally, a request to convert in parallel, compile, and distribute content is received and processed by pipeline application 122 running at server 120. The request may include a set of deliverables, each deliverable including a content item and a corresponding format to which to convert the content item. The request may also include instructions to compile the deliverables and distribute them over selected delivery channels. Pipeline application 122 logs the request and prepares each content item for conversion. Pipeline application 122 then divides processing into a plurality of concurrent processing “threads” 122a-d, each thread corresponding to an associated deliverable. Each production thread converts its associated content item to its selected conversion format. Pipeline application 122 then compiles the converted deliverables from each processing thread 122a-d and distributes the compilation to presentation clients 130a-n for presentation.

[0033] At step 410, a request is received at server 120. The request may be made at request client 210 and submitted to server 120 via communications network 140. User interface 210a at request client 210 may be provided to enable a participating user to make the request.

[0034] The request may include a set deliverables. Each such deliverable may include a selected content item and a selected file format to which to convert the selected content item. For example, a set of deliverables may include a “Mathematics Test” and an “English Test to be converted from XML format into PDF format and into WAV format.

[0035] The request may also include customization requests for each deliverable. Customization requests may be included to adjust the content prior to its conversion and delivery. For example, the request may include a time window in which content or portions of content may be presented at presentation clients 130a-n. The time window may me defined by a start timestamp and a stop timestamp. The timestamp may be a value maintained by server 120. Such a time window is a security feature which may prevent unauthorized presentation of the content at an improper time. The request may also include a language such as, for example, Spanish or French, to which to translate content. The request may also include content objects such as, for example, pictures, video content, and audio content to be converted inserted into each deliverable. The request may also require authorization such as for example, a valid password or identifier to retrieve and present each deliverable at presentation clients 130a-n.



[0036] At step 412, pipeline application 122 logs the request. Step 412 is an optional step. Each deliverable within the request may be logged and cued for processing. Server 120 may employ a cueing service to manage the cueing of requests. Each received request may be stored in a request history maintained by a cueing / tracker database 112. Each request entry in the history may include a corresponding identifier and information such as, for example, a request start timestamp, a current status, and a request complete timestamp. The history enables a user to determine the status of all requests processed by pipeline application 122.

[0037] At step 414, pipeline application 210a prepares the requested content for conversion. Preparing the content may include, for example, retrieving the requested content and customizing the requested content. Content may be retrieved from content database 110 via communications network 140. As should be appreciated, multiple databases or other content storage sources may be used in accordance with the present invention.

[0038] Content may be customized to adjust the content prior to its conversion and delivery. Customization of the content may include, for example, verifying the validity of the content by, for example, verifying that the requested content is present in content database 110 in the appropriate format. Customization of content may also include verifying the user's authorization to access the content, inserting a tag specifying a time window in which the content may be presented at presentation clients 130a-n, translating the content to a language such as, for example, Spanish or French, or otherwise modifying the content in accordance with an instruction from the user.

[0039] At step 416, pipeline application 210a converts each deliverable to its requested file format. The processing steps required to convert each deliverable are format specific and may be dependent on customizations selected by the user. Consequently, at step 418, pipeline application 122 divides the conversion of deliverables into a plurality of concurrent processing threads 122a-d. Each such thread 122a-d corresponds to the conversion of an associated deliverable. For example, thread 122a may process conversion of the "Mathematics Test" from an XML format into a PDF format. Thread 122b may process the conversion of the "Mathematics Test" from an XML format into WAV format. Thread 122c may process conversion of the "English Test" from an XML format into a PDF format. Thread 122b may process the conversion of the "English Test" from an XML format into WAV format.

[0040] A cross sectional view of pipeline application 122 at conversion step 416 in accordance with the present invention is shown in Fig. 5. Pipeline 122 is divided into threads

1220a-d, each corresponding to an associated deliverable. As should be appreciated, pipeline application 122 is not limited to four processing threads and may be divided into any number of processing threads corresponding to any number of selected deliverables. The division of processing into concurrent threads 122a-d enables processing steps specific to the conversion of each deliverable to be performed concurrently without the need to serially complete the conversion of a first deliverable before the conversion of a second deliverable may be commenced.

[0041] A flowchart of an exemplary method for content conversion in accordance with the present invention is shown in Fig. 6. The processing steps shown in Fig. 6 are described in connection with thread 122a, corresponding to the conversion of the “Mathematics test” from an XML file into a PDF file. However, the method shown in Fig. 6 is not limited to converting XML content to PDF content and may be applied to conversion of other such content into other such formats by modifying steps 610-614 in accordance with the particular content, selected conversion format, and customization requests associated with the corresponding deliverable.

[0042] At step 610, thread 122a parses the “Mathematics Test” to retrieve information such as, for example, elements and attributes necessary to convert the test from XML to PDF format. Thread 122a may employ a commercially available plug-in such as, for example, an XPP plug-in to parse the XML file.

[0043] At step 612, the parsed XML content may be converted into a page description (PDL) language such as, for example, PostScript. Such a PDL may be an object oriented language that treats images including fonts as collections of geometrical objects rather than as bit maps. Preferably, the size of such images may be changed using PDL commands.

[0044] At step 614, the PDL content is converted to a file format such as, for example, PDF, which is well suited for distributing and presenting a “published” or printed version of the content. Once converted to PDF, the content may be distributed to presentation clients 130a-n and displayed on an attached monitor or printed using an attached printing device. Acrobat Reader is a reading application developed by Adobe Systems of San Jose, California, which is required at presentation clients 130a-n to present PDF content.

[0045] Referring back to Fig. 4, at step 418, the content converted by threads 120a-d is compiled. Such converted content is preferably compiled into a single compilation file such as, for example, a tape archive (TAR) file. TAR is a file format well known in the art of content

provision that is suitable for distribution of compiled content over communications network 140. For example, a plurality of PDF and WAV files may be compiled into a single TAR file.

**[0046]** At step 420, the compilation may be distributed to the delivery channels selected from delivery channels drop down menu 340. Once distributed, the compiled deliverables may be presented at presentation clients 130a-n. Presentation clients 130a-n may be used in connection with attached presentation devices such as, for example, monitors, speakers, or printers. Such monitors may be used, for example, for presenting computer based tests. Such speakers may be used, for example, for presenting interactive voice response (IVR) tests. Such printers may be used, for example, for presenting printed “pencil and paper” tests. Presentation clients 130a-n may also be used in connection with attached input devices such as, for example, keyboards, mice, or microphones, which enable interaction with the presented content.

**[0047]** Presentation clients 130a-n may have access to a timestamp maintained by server 120 to ensure that distributed content is presented within a desired time window. Presentation clients 130a-n may also require a valid password or identifier to access the distributed content.

**[0048]** While the present invention has been described in connection with the preferred embodiments of the various figures, it is to be understood that other similar embodiments may be used or modifications and additions may be made to the described embodiment for performing the same function of the present invention without deviating therefrom. Therefore, the present invention should not be limited to any single embodiment, but rather should be construed in breadth and scope in accordance with the appended claims.